UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
10/660,242	09/11/2003	Dharmendra Shantilal Modha	AM9990184US2	5338

29154 7590 03/29/2007

FREDERICK W. GIBB, III
Gibb & Rahman, LLC
2568-A RIVA ROAD
SUITE 304
ANNAPOLIS, MD 21401

EXAMINER	
VEILLARD, JACQUES	

ART UNIT	PAPER NUMBER
2165	

MAIL DATE	DELIVERY MODE
03/29/2007	PAPER

Please find below and/or attached an Office communication concerning this application or proceeding.

SUPPLEMENTAL Notice of Allowability	Application No. 10/660,242	Applicant(s) MODHA ET AL.	
	Examiner Jacques Veillard	Art Unit 2165	

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address--

All claims being allowable, PROSECUTION ON THE MERITS IS (OR REMAINS) CLOSED in this application. If not included herewith (or previously mailed), a Notice of Allowance (PTOL-85) or other appropriate communication will be mailed in due course. **THIS NOTICE OF ALLOWABILITY IS NOT A GRANT OF PATENT RIGHTS.** This application is subject to withdrawal from issue at the initiative of the Office or upon petition by the applicant. See 37 CFR 1.313 and MPEP 1308.

1. ☒ This communication is responsive to *5/4/2006*.

2. ☒ The allowed claim(s) is/are *28-40, 55-61 (Renumbered claims 1-20)*.

3. ☐ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).
   a) ☐ All b) ☐ Some* c) ☐ None of the:
      1. ☐ Certified copies of the priority documents have been received.
      2. ☐ Certified copies of the priority documents have been received in Application No. _____ .
      3. ☐ Copies of the certified copies of the priority documents have been received in this national stage application from the International Bureau (PCT Rule 17.2(a)).

   * Certified copies not received: _____.

Applicant has THREE MONTHS FROM THE "MAILING DATE" of this communication to file a reply complying with the requirements noted below. Failure to timely comply will result in ABANDONMENT of this application.
THIS THREE-MONTH PERIOD IS NOT EXTENDABLE.

4. ☐ A SUBSTITUTE OATH OR DECLARATION must be submitted. Note the attached EXAMINER'S AMENDMENT or NOTICE OF INFORMAL PATENT APPLICATION (PTO-152) which gives reason(s) why the oath or declaration is deficient.

5. ☐ CORRECTED DRAWINGS (as "replacement sheets") must be submitted.
   (a) ☐ including changes required by the Notice of Draftsperson's Patent Drawing Review (PTO-948) attached
      1) ☐ hereto or 2) ☐ to Paper No./Mail Date _____.
   (b) ☐ including changes required by the attached Examiner's Amendment / Comment or in the Office action of Paper No./Mail Date _____.

   Identifying indicia such as the application number (see 37 CFR 1.84(c)) should be written on the drawings in the front (not the back) of each sheet. Replacement sheet(s) should be labeled as such in the header according to 37 CFR 1.121(d).

6. ☐ DEPOSIT OF and/or INFORMATION about the deposit of BIOLOGICAL MATERIAL must be submitted. Note the attached Examiner's comment regarding REQUIREMENT FOR THE DEPOSIT OF BIOLOGICAL MATERIAL.

Attachment(s)

1. ☐ Notice of References Cited (PTO-892)
2. ☐ Notice of Draftperson's Patent Drawing Review (PTO-948)
3. ☐ Information Disclosure Statements (PTO/SB/08), Paper No./Mail Date _____
4. ☐ Examiner's Comment Regarding Requirement for Deposit of Biological Material
5. ☐ Notice of Informal Patent Application
6. ☒ Interview Summary (PTO-413), Paper No./Mail Date *3/27/2007* .
7. ☒ Examiner's Amendment/Comment
8. ☐ Examiner's Statement of Reasons for Allowance
9. ☐ Other _____.

Jacques Veillard

HOSAIN ALAM
SUPERVISORY PATENT EXAMINER

DETAILED ACTION

EXAMINER'S AMENDMENT

1. An examiner's amendment to the record appears below. Should the changes and/or additions be unacceptable to applicant, an amendment may be filed as provided by 37 CFR 1.312. To ensure consideration of such an amendment, it MUST be submitted no later than the payment of the issue fee.

Authorization for this examiner's amendment was given in a telephone interview with attorney Mohammad S. Rahman (Reg. No. 43,029) the undersigned for Applicant(s) on 3/27/07.

The application has been amended as follows:

In the Specification;

Please replace pages 18, 26, and 27 in the original specification by the new attached pages 18, 26, and 27 in order to be consistent with equation number 6 on page 18, the last two equation on page 26, and equation number 8 on page 27. See the attached papers.

Conclusion

2. Any inquiry concerning this communication or earlier communications from the examiner should be directed to Jacques Veillard whose telephone number is (571) 272-4086. The examiner can normally be reached on Mon. to Fri. from 9 AM to 4:30 PM, alt. Fri. off..

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, Jeffrey Gaffin can be reached on (571) 272- 4146. The fax phone number for the organization where this application or proceeding is assigned is 571-273-8300.

Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Status information for published applications may be obtained from either Private PAIR or Public PAIR. Status information for unpublished applications is available through Private PAIR only. For more information about the PAIR system, see http://pair-direct.uspto.gov. Should you have questions on access to the Private PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free). If you would like assistance from a USPTO Customer Service Representative or access to the automated information system, call 800-786-9199 (IN USA OR CANADA) or 571-272-1000.

J.V.
J.V

Jacques Veillard
Patent Examiner TC 2100

March 27, 2007

Intuitively, the objective function measures the combined coherence of all the k clusters.

The Process. The invention's objective is to find ***k*** disjoint clusters

$\pi_1^{\dagger}, \pi_2^{\dagger}, \ldots\ldots, \pi_3^{\dagger}$ such that the following is maximized

$$\left\{\pi_j^{\dagger}\right\}_{j=1}^{k} = \frac{\arg \max}{\left\{\pi_j\right\}_{j=1}^{k}} \left(\sum_{j=1}^{k} \sum_{x \pi_j} S(x, c_j) \right) \qquad (6)$$

Even when only one of the parameters α_d, α_f, or α_b is nonzero, finding the optimal solution to the above maximization problem is known to be **NP**-complete. Discussed below is an efficient and effective approximation process: the *toric k-means* that may be thought of as a *gradient ascent* method.

Step 1. Start with an arbitrary partitioning of the document vectors, namely, $\left\{\pi_j^{(0)}\right\}_{j=1}^{k}$. Let $\left\{c_j^{(0)}\right\}_{j=1}^{k}$ denote the concept triplets associated with the given partitioning. Set the index of iteration $t = 0$. The choice of the initial partitioning is quite crucial to finding a "good" local minima; for recent work on this area, see (Bradley, P., and Fayyad, U. Refining initial points for k-means clustering In *ICML* (1998), pp. 91-99, incorporated herein by reference).

Step 2. For each document vector triplet x_i, $1 < i < n$ the invention finds the concept triplet that is closest to x_i. Now, for $1 < j < k$, compute the new partitioning $x_i, 1 \le i \le n$

Throughout this section, the invention fixes the number of clusters k $\geq$2. As before, let α_d , α_f , and α_b be nonnegative numbers that sum to 1. Geometrically, these parameters lie on a planar triangular region, say, Δo, that is shown in Figure 5. For brevity, the invention writes $\alpha = (\alpha_d , \alpha_f , \alpha_b)$. Let $\Pi(\alpha) = \{\pi_j\}_{j=1}^{k}$, denote the partitioning obtained by running the toric *k*-means process with the parameter values α_d, α_f, and α_b. From the set of all possible clusterings $\{\Pi(\alpha) : \alpha \in \Delta_0\}$. The invention selects a partitioning that yields the *best* cluster annotations. Towards this goal, the invention introduces a figure-of-merit for evaluating and comparing various clusterings.

Fixing a clustering $\Pi(\alpha)$. For the given clustering, the summary, which is a descriptive characteristic, for each of the clusters will be good if each cluster is as coherent as possible in the word feature space, that is, if the following is maximized:

$$\Gamma_d(\alpha) \equiv \Gamma_d(\Pi(\alpha)) = \sum_{j=1}^{k} \sum_{X \in \pi_j} D^T D_j^*$$

where $\boldsymbol{x} = (\boldsymbol{D}, \boldsymbol{F}, \boldsymbol{B})$. Furthermore, the keywords, which are a discriminative characteristics, will be good if the following is minimized:

$$\Lambda_d(\alpha) \equiv \Lambda_d(\Pi(\alpha)) = \sum_{j=1}^{k} \sum_{X \in \pi_j} \sum_{\ell=1, \ell \neq j}^{k} D_j^T D_\ell^*,$$

where $\boldsymbol{x} = (\boldsymbol{D}, \boldsymbol{F}, \boldsymbol{B})$. Intuitively, $\Gamma_d(\alpha)$ and $\Lambda_d(\alpha)$ capture the *average within cluster coherence and average between cluster coherence,* respectively, of the

clustering $\Pi(\alpha)$ in the word feature space. The summary and the keywords both will be good if the following ratio is maximized:

$$Qd(a) \equiv Q_d(II(a)) = \begin{cases} \left(\frac{\Gamma_d(a)}{\Lambda_{d(a)}}\right)^{nd/n} & \text{if } \Lambda_d(\alpha) > 0, \\ 1 & \text{if } \Lambda_d(\alpha)=0, \end{cases} \tag{8}$$

where n_d denotes the number of document triplets in $\boldsymbol{Q}$ that have a non-zero word feature vector; see, for example, Figure 3. In the case that $\Lambda_d(\alpha) = 0$, the clusters are *perfectly separated* in the word feature space.

The quantities $\Gamma_f(\alpha)$, $\Lambda_f(\alpha)$, $\Gamma_b(\alpha)$, $\Lambda_b(\alpha)$, $Q_f(\alpha)$, and $Q_b(\alpha)$ are defined in a similar fashion. The quantity $Q_f(\alpha)$should be maximized to obtain good quality review and references, and the Quantity $Q_b(\alpha)$ should be maximized to obtain good quality breakthrough and citations.

In Figure 5 the triangular region Δ_o is formed by the intersection of the plane $\alpha_d + \alpha_f + \alpha_b = 1$ with the nonnegative orthant of R^3. The left-vertex, the right-vertex, and the top-vertex of the triangle correspond to the points (1, 0, 0), (0, 1, 0), and (0, 0, 1), respectively.

The following shows how the invention selects the optimal parameter tuple $\alpha^\dagger$ and the corresponding clustering $II(\alpha^\dagger)$.

Step 1. Theoretically, the invention would like to run the tortic k-means process for every parameter triplet in:

$$\Delta_0 = \{\alpha : \alpha_d + \alpha_f + \alpha_b = 1, \alpha_d, \alpha_f, \alpha_b \geq 0\}. \tag{9}$$